

On the Influence of Bias and of Personal Equation in Statistics of Ill-Defined Qualities: An Experimental Study

G. Udny Yule

Proc. R. Soc. Lond. A 1906 **77**, 337-339

doi: 10.1098/rspa.1906.0030

Email alerting service

Receive free email alerts when new articles cite this article - sign up in the box at the top right-hand corner of the article or click [here](#)

To subscribe to *Proc. R. Soc. Lond. A* go to: <http://rspa.royalsocietypublishing.org/subscriptions>

On the Influence of Bias and of Personal Equation in Statistics of Ill-defined Qualities: An Experimental Study.

By G. UDNY YULE, Newmarch Lecturer in Statistics, University College, London.

(Communicated by Professor O. Henrici, F.R.S. Received November 4,—Read December 7, 1905.)

(Abstract.)

I. This experiment was undertaken to elucidate the real character of such statistics as those of eye-colour, hair-colour, temper, health, etc., which have been given, *e.g.*, by Mr. Galton and by Professor Pearson. The statistics are, it should be noted, not merely statistics of *qualities*, but of *ill-defined qualities*, the only guidance to the use of the terms of classification being—with some exceptions—common usage. Strictly speaking we must remember that data so collected are statistics, not of qualities themselves, but of names assigned thereto. It was desired to determine how far the distinction is of importance (1) as regards the naming of single samples: (2) as regards the naming of pairs, two samples of a quality being named more or less together, by themselves, for forming a contingency-table.

A matt-surfaced photographic paper was printed by successive exposure to 16 depths of tint, from a slightly impure white to a deep blackish-brown. Small scraps of about $\frac{3}{8}$ -inch square were cut from each tint, and mounted on cards, two scraps being placed on each card, combined in such a way that every possible combination occurred, making $16 \times 16 = 256$ cards. Observers were then asked to name the tints on each card under one or more of the following schemes of classification, each observer naming the whole pack:—

Series A.—1. Light. 2. Dark.

Series B.—1. Light. 2. Medium. 3. Dark.

Series C.—1. Very light to light. 2. Rather light. 3. Medium. 4. Rather dark. 5. Dark to very dark.

The cards in the pack were arranged, by shuffling, in a more or less random order. Returns were obtained from 34 volunteer observers, who sent in 17 schedules under Series A, 20 under Series B, and 30 under Series C.

II. As regards the way in which single tints alone are named: (1) No observer, as might be expected, is quite self-consistent in his naming; (2) the inconsistencies are greater for Series B than for A, and greater for C than for B; (3) the observers attach very sensibly different meanings to the terms used for classification; (4) As a combined result of (1) and (3) the terms used for classification do not determine discrete classes, but

338 Mr. G. Udny Yule. *Influence of Bias and of* [Nov. 4,

very widely overlapping frequency distributions. In Series B, for example, the following is the distribution of the actual tints under the names :—

Table showing Distribution of Tints under Names for the whole of the Twenty Observers who made Returns under Series B. 10,240 Observations.

Name assigned to tint.	Number of tint (1 = white, 16 = very dark brown).																Total.
	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.	
Light ...	640	640	636	618	521	368	269	186	93	23	15	—	—	—	—	—	4009
Medium ...	—	—	4	22	119	272	366	451	524	564	516	227	193	21	6	1	3286
Dark ...	—	—	—	—	—	—	5	3	23	53	109	413	447	619	634	639	2945

The result suggests that methods which treat as discrete classes determined only by names in common use are not strictly applicable, and that quantitative results obtained by such methods can be regarded only as useful illustrative analogies.

III. Contingency-tables were formed from each observer's schedules for the names assigned to the upper and lower tints on each card. If an observer returned quite without bias, the frequencies in the compartments of his table should be given by the rule of independence (total of row $\times$ total of column $\div$ by whole number of observations). There proved, however, to be a distinct tendency to return an excess of pairs of the same name; this tendency, though vanishingly small for Series A, became marked for Series B, and more marked still for Series C. This feature was emphasised when different observers' results were pooled, as the pooling of results of different observers *who are quite unbiassed*, tends in itself to give an excess of homonymous pairs. In Series C there was also an excess of contrasted pairs. The following table gives the actual aggregate of returns for Series B. The first number is the theoretically correct frequency, the number after the sign the excess or deficiency of the actual returns.

Name assigned to lower tint.	Name assigned to upper tint on card.		
	Light.	Medium.	Dark.
Light	785 + 65	633 - 62	583 - 3
Medium	653 - 35	527 + 66	486 - 31
Dark	570 - 30	460 - 4	423 + 34

The above table includes returns from the 20 observers ; separate tables for the first and second ten in alphabetical order gave reasonably consistent results, the frequencies of homonymous pairs being in excess in every case.

An experiment was also tried, eliminating the returns for certain cards in the pack with contrasted tints so as to make the distribution a correlated instead of an independent one. In this case the contrasted pairs returned for Series B seem to be in excess instead of in defect as above, the correlation coefficient calculated for a division between light and medium for the one tint and medium and dark for the other, being slightly lower than the true value for a similar division of the actual data. For symmetrical division the coefficient was slightly higher than the true.

IV. Certain of the contingency-tables given by Professor Pearson were examined to see how far the observed peculiarities might be due to subjective influences. The excess of homonymous pairs, and, indeed, the correlation, in the eye-colour table, for homogamy between husband and wife would seem to be largely due to such influences. In eye-colour tables for brother and brother, and for father and son, the divergencies of the distribution from normality are of the same type as the divergencies, in the above experiment, of observation from truth, but much larger. For the brother-brother tables for temper and for curliness of hair, speaking generally, the same thing holds. There is very possibly, accordingly, some real objective effect, the nature of which requires elucidating, but considerable reserve is necessary in view of the qualitative similarity to effects of subjective origin. A collection of data is required in which these are eliminated by the use of good representative scales, or possibly by the naming of the two members of a pair quite independently by different observers.
